

Projected Hartree Fock Theory as a Polynomial Similarity Transformation Theory of Single Excitations

Yiheng Qiu

Department of Chemistry, Rice University, Houston, TX 77005-1892

Thomas M. Henderson and Gustavo E. Scuseria

Department of Chemistry, Rice University, Houston, TX 77005-1892 and
Department of Physics and Astronomy, Rice University, Houston, TX 77005-1892

(Dated: August 2, 2016)

Spin-projected Hartree-Fock is introduced as a particle-hole excitation ansatz over a symmetry-adapted reference determinant. Remarkably, this expansion has an analytic expression that we were able to decipher. While the form of the polynomial expansion is universal, the excitation amplitudes need to be optimized. This is equivalent to the optimization of orbitals in the conventional projected Hartree-Fock framework of non-orthogonal determinants. Using the inverse of the particle-hole expansion, we similarity transform the Hamiltonian in a coupled-cluster style theory. The left eigenvector of the non-hermitian Hamiltonian is constructed in a similar particle-hole expansion fashion, and we show that to numerically reproduce variational projected Hartree-Fock results, one needs as many pair excitations in the bra as the number of strongly correlated entangled pairs in the system. This single-excitation polynomial similarity transformation theory is an alternative to our recently presented double excitation theory, but supports projected Hartree-Fock and coupled cluster simultaneously rather than interpolating between them.

The main difficulty in computational quantum chemistry is the need for an accurate description of electronic correlation effects. When electrons are weakly correlated and a mean-field picture is qualitatively accurate, it is probably fair to say that this difficulty is overcome by using some form of coupled cluster (CC) theory.¹⁻⁴ When the mean-field picture is qualitatively incorrect, however, and electrons are strongly correlated, the situation is significantly different, and it is probably equally fair to say that there is no completely general solution to the problem.⁵ When the number of strongly correlated electrons is not too large, active-space methods work well, but ultimately these become too computationally cumbersome to be of practical utility.

One appealing approach to the strong correlation problem is the use of projected Hartree-Fock (PHF).⁶⁻¹¹ When a system becomes strongly correlated, the mean-field tends to signal its own demise by spontaneously breaking a symmetry of the Hamiltonian. The component of the broken-symmetry mean-field wave function which has the correct symmetries will be a multi-determinantal wave function which typically offers a fairly reliable description of the strong correlations. Moreover, by optimizing the mean-field in the presence of the symmetry projection operator, one can *deliberately* break symmetries even when those symmetries do not break spontaneously; this variation after projection approach to symmetry breaking and restoration leads to wave functions which are well-behaved as a function of Hamiltonian parameters.^{10,11}

To see all this in action, we show in Fig. 1 a plot of the dissociation of the N_2 molecule in a minimal basis, where exact results are available. The key features we wish to emphasize are as follows. First, the symmetry-adapted mean-field (RHF) is not terrible near equilib-

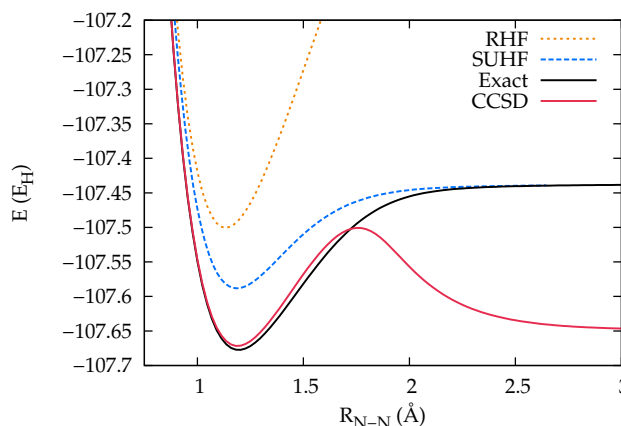


FIG. 1. Dissociation of N_2 in the STO-3G basis.

rium, but is useless toward dissociation. Second, coupled cluster with single and double excitations (CCSD) based on the symmetry-adapted reference is highly accurate near equilibrium where RHF is reasonable, but as the bond stretches, RHF breaks down, and the molecule becomes more strongly correlated, CCSD goes haywire. Finally, spin-projected unrestricted Hartree-Fock (SUHF) is well-behaved everywhere but is clearly missing a chunk of the correlation energy – that is, it lacks an accurate accounting for the weak correlations that CCSD so readily recovers. We note in passing that the exactness of SUHF at dissociation is a consequence of the minimal basis set and is not a general result.

The breakdown may be even clearer in the Hubbard Hamiltonian,¹² as shown in Fig. 2. Here, we see that for small on-site repulsion U where the system is weakly

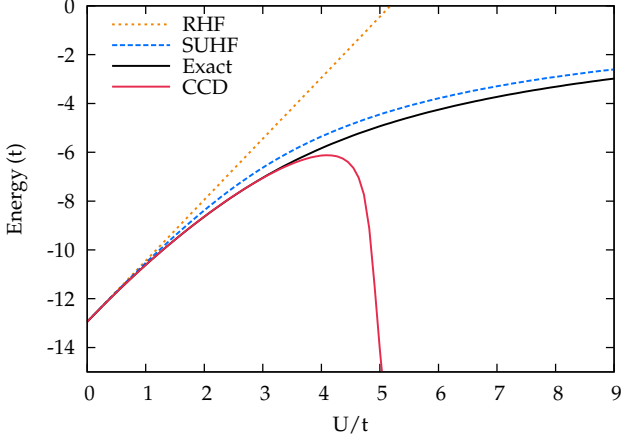


FIG. 2. Total energies in the 10-site half-filled periodic one-dimensional Hubbard Hamiltonian.

correlated, CCSD is exceptionally accurate, but that it breaks down completely for larger U ; SUHF, meanwhile, is reasonable but imperfect everywhere.^{13,14}

In an ideal world, one could simply combine PHF and coupled cluster, but this task is complicated by the very different natures of the two theories. The PHF wave function is concisely written as a (short) expansion in a set of non-orthogonal broken-symmetry determinants, while the coupled cluster wave function is written as a (long) expansion in a set of orthogonal symmetry-adapted determinants. The PHF energy is an expectation value, and the PHF wave function is obtained using the variational principle, where coupled cluster uses a projective Schrödinger equation approach to define a non-variational energy. It is difficult to see how to cleanly marry these two approaches.

In this communication, building on our work on polynomial similarity transformations (PoST),¹⁵ we show how one can cast PHF in terms of particle-hole excitations out of a symmetry-adapted determinant and, moreover, we show how one can optimize the PHF wave function in coupled-cluster-like fashion as opposed to the more traditional variational optimization. We will here consider only what we call SUHF, and will further limit ourselves to projection onto singlet states ($s = 0$) only, but the theory can be extended to a much more general framework including other quantum numbers and symmetries.

We start from the observation that an $m_S = 0$ unrestricted Hartree-Fock (UHF) determinant can be written as a Thouless transformation of an RHF determinant:

$$|\text{UHF}\rangle = e^{T_1 + U_1} |\text{RHF}\rangle, \quad (1)$$

where we have introduced

$$T_1 = \sum_i t_i^a E_a^i, \quad (2a)$$

$$U_1 = \sum_i u_i^a S_a^i, \quad (2b)$$

$$E_a^i = c_{a\uparrow}^\dagger c_{i\uparrow} + c_{a\downarrow}^\dagger c_{i\downarrow}, \quad (2c)$$

$$S_a^i = c_{a\uparrow}^\dagger c_{i\uparrow} - c_{a\downarrow}^\dagger c_{i\downarrow}. \quad (2d)$$

We follow the convention that spatial orbitals indexed $i, j, k, \dots (a, b, c, \dots)$ are occupied (empty) in $|\text{RHF}\rangle$. Note that the Thouless transformation above does not preserve normalization and that we have assumed that $|\text{UHF}\rangle$ and $|\text{RHF}\rangle$ are not orthogonal.

A key result in this paper is that the SUHF wave function, traditionally obtained via a singlet symmetry projection operator P acting on $|\text{UHF}\rangle$, can be written as

$$|\text{SUHF}\rangle = P |\text{UHF}\rangle = e^{T_1} F(K_2) |\text{RHF}\rangle, \quad (3)$$

where K_2 is the singlet component of U_1^2 ,

$$K_2 = -\frac{1}{6} \sum (u_i^a u_j^b + 2 u_i^b u_j^a) E_a^i E_b^j, \quad (4)$$

and the polynomial $F(K_2)$ is given by

$$F(K_2) = \frac{\sinh(\sqrt{6} K_2)}{\sqrt{6} K_2} \quad (5a)$$

$$= 1 + K_2 + \frac{3}{10} K_2^2 + \frac{3}{70} K_2^3 + \dots \quad (5b)$$

Only even powers of U_1 appear, as odd powers of U_1 break spin symmetry, a fact already noted in Ref. 16, although these authors did not point out that the PHF wave function could be expressed as a polynomial of the double-excitation operator K_2 , a key result needed for solving for the u_i^a amplitudes as we do here.

Details of the proof will be presented elsewhere, but a rough sketch of the idea proceeds as follows. The expression for each term $P U_1^n$ can be obtained by analytically integrating over spin rotation angles, in a manner basically analogous to the numerical integration done in SUHF among non-orthogonal determinants, but working here with orthogonal particle-hole excitations. That $F(K_2)$ is given by the sinh polynomial is recognized by direct inspection of the individual projected terms and numerically proven below by comparison with our previous implementation of SUHF.¹¹ Optimization of the coefficients u_i^a plays the role of orbital optimization in SUHF.¹¹

Thus far, all we have done is to reparameterize the SUHF wave function. This is an important step, but it should be noted that conventionally SUHF defines the energy as an expectation value which it variationally minimizes with respect to the UHF determinant. In our language, this would require us to solve

$$E = \frac{\langle \text{RHF} | F(K_2^\dagger) e^{T_1^\dagger} H e^{T_1} F(K_2) | \text{RHF} \rangle}{\langle \text{RHF} | F(K_2^\dagger) e^{T_1^\dagger} e^{T_1} F(K_2) | \text{RHF} \rangle}, \quad (6a)$$

$$0 = \frac{\partial E}{\partial t_i^a} = \frac{\partial E}{\partial u_i^a}, \quad (6b)$$

which is not readily compatible with the typical approach used in traditional CC theory.

In standard CC theory, we construct a similarity-transformed Hamiltonian

$$\bar{H} = e^{-T} H e^T, \quad (7)$$

where T creates excitations out of the reference, which we will denote simply as

$$T = \sum t_\mu Q_\mu^\dagger \quad (8)$$

where t_μ are excitation amplitudes and excited determinants $|\Phi_\mu\rangle$ are created by the action of Q_μ^\dagger on the reference. We wish to choose the similarity transformation such that $|\text{RHF}\rangle$ is a right-hand eigenstate of \bar{H} . Because \bar{H} is non-Hermitian, we must also solve for a left-hand eigenstate, which is parameterized as $\langle\text{RHF}|(1+Z)$, where

$$Z = \sum z_\mu Q_\mu. \quad (9)$$

We define the energy by the expectation value of \bar{H} , which is made stationary with respect to the amplitudes defining T and Z :

$$E = \langle\text{RHF}|(1+Z) \bar{H}|\text{RHF}\rangle, \quad (10a)$$

$$0 = \frac{\partial E}{\partial t_\mu} = \frac{\partial E}{\partial z_\mu}. \quad (10b)$$

Our goal is thus to write PHF in a coupled-cluster-like language, because if we can do so, then combining CC and PHF is essentially straightforward. To write PHF in this manner, we proceed in analogy with CC above and define a similarity-transformation for PHF:

$$\bar{H}_{\text{PHF}} = F^{-1}(K_2) e^{-T_1} H e^{T_1} F(K_2). \quad (11)$$

Because T_1 and K_2 are both excitation operators, they commute, so we need not worry about the order of $F(K_2)$ and $\exp(T_1)$. Whereas the singlet projection operator P does not have an inverse, the inverse polynomial $F^{-1}(K_2)$ can be extracted from a Taylor series expansion of $1/F(x)$ and begins

$$F^{-1}(K_2) = 1 - K_2 + \frac{7}{10} K_2^2 - \frac{31}{70} K_2^3 + \dots \quad (12)$$

As with CC theory, we want the right-hand eigenstate of \bar{H}_{PHF} to be the symmetry-adapted determinant, but we will need a more complicated left-hand eigenstate to approximate the variational PHF case. Thus, we define

$$E = \langle\text{RHF}|(1+Z_1) G(L_2) \bar{H}_{\text{PHF}}|\text{RHF}\rangle, \quad (13a)$$

$$Z_1 = \sum z_a^i E_a^i, \quad (13b)$$

$$L_2 = -\frac{1}{6} \sum \left(v_a^i v_b^j + 2 v_b^i v_a^j \right) E_a^i E_b^j, \quad (13c)$$

$$E_i^a = (E_i^a)^\dagger = c_{i\uparrow}^\dagger c_{a\uparrow} + c_{i\downarrow}^\dagger c_{a\downarrow}. \quad (13d)$$

where $G(L_2)$ is a polynomial we will specify shortly. We will make the energy stationary with respect to the four distinct single-excitation amplitudes:

$$0 = \frac{\partial E}{\partial t_i^a} = \frac{\partial E}{\partial u_i^a} = \frac{\partial E}{\partial z_a^i} = \frac{\partial E}{\partial v_i^a}. \quad (14)$$

Note the strong resemblance to coupled-cluster doubles, which has

$$T = T_2 = \frac{1}{2} \sum t_{ij}^{ab} E_a^i E_b^j, \quad (15a)$$

$$Z = Z_2 = \frac{1}{2} \sum z_{ab}^{ij} E_i^a E_j^b, \quad (15b)$$

though here we have a more complicated left-hand state and, crucially, K_2 and L_2 have factorizable amplitudes while T_2 and Z_2 in general do not.

There are two general strategies we might pursue for the bra polynomial $G(L_2)$. One is to make $G(x) = F(x)$ in analogy with extended coupled cluster theory.^{17,18} The second is to try to match this projective PHF energy expression to the variational one, in which case we want to have schematically

$$G(L_2) F^{-1}(K_2) \approx \frac{F(K_2^\dagger)}{\langle\text{SUHF}|\text{SUHF}\rangle}. \quad (16)$$

While we cannot in general enforce this condition exactly, we *can* enforce it on average. In other words, we can adjust the coefficients c_n in $G(L_2) = \sum c_n L_2^n$ by imposing conditions like

$$\langle\text{RHF}|G(L_2) F^{-1}(K_2)|n\rangle = \frac{\langle\text{RHF}|F(K_2^\dagger)|n\rangle}{\langle\text{SUHF}|\text{SUHF}\rangle}, \quad (17)$$

where $|n\rangle$ stands for n -tuply excited determinants. This leads to a set of linear equations for the coefficients c_n . We consider both approaches.

We should say a few words about computational complexity. For connected (non-factorizable) double-excitation operators, evaluating either expectation values of $F(K_2^\dagger) H F(K_2)$ or $G(L_2) \bar{H}_{\text{PHF}}$ would be prohibitively expensive unless we truncated the polynomials $F(x)$ and $G(x)$ to low order, but a key difference between our approach and CC theory is that our energy expression does not truncate; only similarity transformations generated by exponentials lead to terminating series (at $\mathcal{O}(T^4)$ for a two-body Hamiltonian). However, as we shall discuss in detail in a subsequent manuscript, the evaluation of terms even with high powers of K_2 or L_2 is feasible, essentially because K_2 and L_2 factorize, and yields relatively simple expressions with polynomial computational cost. Nevertheless, we shall here show results where we manually truncate $G(L_2)$, so one can assess convergence toward PHF.

Figure 3 shows energies relative to SUHF for the 14-site Hubbard Hamiltonian. The dot circles represent energies obtained via Eq. (6), numerically demonstrating the exactness of the sinh polynomial of Eq. (5). It is

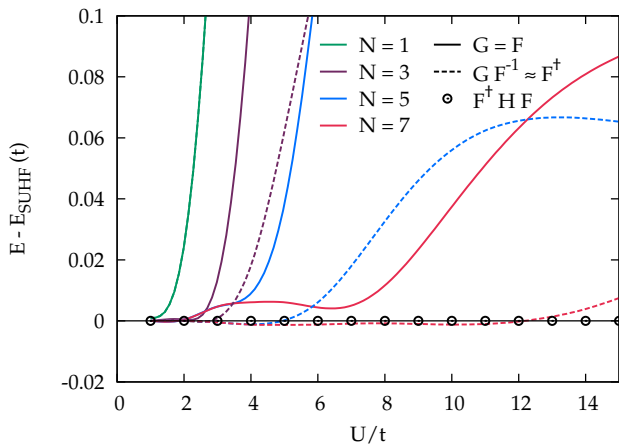


FIG. 3. Energies relative to SUHF in the half-filled 14-site Hubbard Hamiltonian. Solid lines indicate $G(x) = F(x)$ and dashed lines indicate we have fit $G(L_2) F^{-1}(K_2) \sim F(K_2^\dagger)$. The label “N” indicates that $G(x)$ has been truncated at L_2^N . Dots denote the variational energy expression.

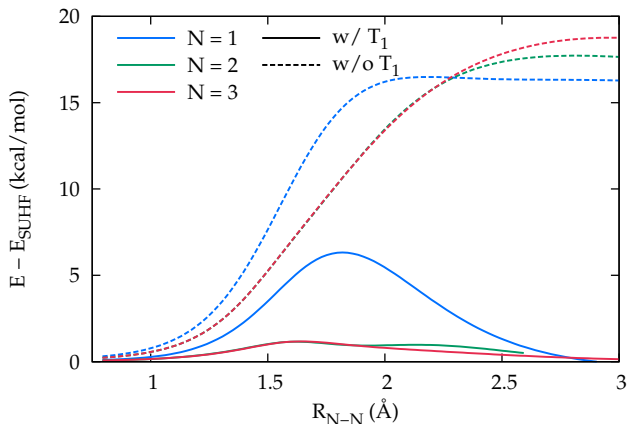


FIG. 4. Energies relative to SUHF in the dissociation of N_2 , using the cc-pVDZ basis. We fit $G F^{-1} \sim F^\dagger$ with G truncated to order N ; solid and dashed lines show results with and without the inclusion of T_1 .

clear that a low-order truncation of $G(L_2)$ is inadequate; although not shown, the same is true for $F(K_2^\dagger)$. Second, with $G = F$, our projective approach replicates standard PHF to a reasonable extent. Finally, when we optimize the polynomial G to match the variational expectation value, we reproduce PHF even better. The remaining small discrepancies between our projective approach and

the exact SUHF are presumably because we cannot make $G(L_2) F^{-1}(K_2)$ match $F(K_2^\dagger)$ exactly. Note that in the Hubbard Hamiltonian, $T_1 = 0$ by symmetry, so these results are obtained purely with K_2 (or, if one prefers, U_1).

We can see the importance of single excitations in Fig. 4 where we show results for the dissociation of N_2 in the cc-pVDZ basis with an optimized polynomial $G(L_2)$. As in the Hubbard Hamiltonian we converge toward SUHF as we increase the degree of the polynomial G , but only when we include T_1 . This is of course what one would expect on the basis of Eq. (1).

We wish to emphasize that while one can truncate $F(K_2^\dagger)$ or $G(L_2)$, this truncation only really converges at order N , where N is the number of strongly correlated pairs. Thus, in the half-filled 14-site Hubbard Hamiltonian where there are 14 strongly correlated electrons, we must in general retain terms all the way up to L_2^7 , where in N_2 where there are six strongly correlated electrons we need keep terms only up to L_2^3 . This is in analogy with the conventional wisdom in standard coupled cluster theory. Unlike in standard coupled cluster, retaining terms of high order is straightforward because K_2 and L_2 are ultimately obtained from *single* excitation amplitudes.

Finally, let us reiterate one last time what we wish to demonstrate in this work. Previous work on polynomial similarity transformations¹⁵ showed how one could interpolate between the number projected BCS and coupled cluster forms of wave function by writing them both in the same language. Rather than interpolating, however, we would like to combine coupled cluster theory and symmetry-projected mean-field methods in a more sophisticated and presumably more correct wave function, a task made more difficult by the very different natures of the two theories (but see Ref. 19 for a possible solution in the broken symmetry basis). In this manuscript, we have shown how to use a coupled-cluster-like formalism to solve for the energy and wave function of projected Hartree-Fock theory. By so doing, we set the stage for the explicit wave function-based combination of PHF and CC, thereby obtaining the best of both worlds.

Acknowledgments: This work was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Computational and Theoretical Chemistry Program under Award No.DE-FG02-09ER16053. G.E.S. is a Welch Foundation Chair (C-0036). We thank Jorge Dukelsky and Matthias Degroote for helpful discussions. Jinmo Zhao’s *pyslata* code was helpful for debugging our energy expression.

¹ F. Coester and H. Kümmel, Nuclear Physics **17**, 477 (1960).

² J. Paldus and X. Z. Li, Adv. Chem. Phys. **110**, 1 (1999).

³ R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291

(2007).

⁴ I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics* (Cambridge University Press, New York, 2009) and references contained therein.

- ⁵ J. P. F. LeBlanc, A. E. Antipov, F. Becca, I. W. Bulik, G. K.-L. Chan, C.-M. Chung, Y. Deng, M. Ferrero, T. M. Henderson, C. A. Jiménez-Hoyos, E. Kozik, X.-W. Liu, A. J. Millis, N. V. Prokof'ev, M. Qin, G. E. Scuseria, H. Shi, B. V. Svistunov, L. F. Tocchio, I. S. Tupitsyn, S. R. White, S. Zhang, B.-X. Zheng, Z. Zhu, and E. Gull, *Phys. Rev. X* **5**, 041041 (2015).
- ⁶ P.-O. Löwdin, *Phys. Rev.* **97**, 1509 (1955).
- ⁷ P. Ring and P. Schuck, *The Nuclear Many-Body Problem* (Springer-Verlag, New York, NY, 1980).
- ⁸ J.-P. Blaizot and G. Ripka, *Quantum Theory of Finite Systems* (The MIT Press, Cambridge, MA, 1985).
- ⁹ K. Schmid, *Prog. Part. Nucl. Phys.* **52**, 565 (2004).
- ¹⁰ G. E. Scuseria, C. A. Jiménez-Hoyos, T. M. Henderson, J. K. Ellis, and K. Samanta, *J. Chem. Phys.* **135**, 124108 (2011).
- ¹¹ C. A. Jiménez-Hoyos, T. M. Henderson, T. Tsuchimochi, and G. E. Scuseria, *J. Chem. Phys.* **136**, 164109 (2012).
- ¹² J. Hubbard, *Proc. Roy. Soc. London* **276**, 238 (1963).
- ¹³ R. Rodríguez-Guzmán, K. W. Schmid, C. A. Jiménez-Hoyos, and G. E. Scuseria, *Phys. Rev. B* **85**, 245130 (2012).
- ¹⁴ R. Rodríguez-Guzmán, C. A. Jiménez-Hoyos, and G. E. Scuseria, *Phys. Rev. B* **90**, 195110 (2014).
- ¹⁵ M. Degroote, T. M. Henderson, J. Zhao, J. Dukelsky, and G. E. Scuseria, *Phys. Rev. B* **93**, 125124 (2016).
- ¹⁶ P. Piecuch, R. Toboła, and J. Paldus, *Phys. Rev. A* **54**, 1210 (1996).
- ¹⁷ J. Arponen, *Ann. Phys.* **151**, 311 (1983).
- ¹⁸ P. Fan and P. Piecuch, *Adv. Quantum Chem.* **51**, 1 (2006).
- ¹⁹ T. Duguet, *J. Phys. G* **42**, 025107 (2015).